

IN THE CLAIMS

1. (Previously Amended) A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

identifying sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and

outputting the identified mutual dependence patterns based on results of the comparisons.

2. (Previously Amended) A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

identifying one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and

outputting the one or more identified patterns based on results of the comparisons;

wherein the identifying step further comprises identifying a set of items in the input data set, which includes at least two subsets of at least one item, as a pattern when the set of items has a conditional probability value computed therefor that is not less than a predetermined threshold value, wherein the conditional probability value is indicative of a probability that both of the at least two subsets of at least one item will occur given that one of the at least two subsets of at least one item has occurred.

3. (Original) The method of claim 2, wherein the predetermined threshold value is between zero and one.

4. (Original) The method of claim 2, wherein the conditional probability value is estimated by the number of occurrences in the input data set of the two subsets divided by the number of occurrences in the data set of one of the two subsets.

5. (Previously Amended) A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

identifying one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and

outputting the one or more identified patterns based on results of the comparisons;

wherein the identifying step further comprises identifying a set of items in the input data set as a pattern when the set of items has a conditional probability value computed for the set of items minus a particular item of the set, given the particular item of the set, that is not less than a predetermined threshold value.

6. (Original) The method of claim 5, wherein the predetermined threshold value is between zero and one.

7. (Previously Amended) The method of claim 2, wherein the input data set comprises transaction data.

8. (Previously Amended) The method of claim 2, wherein the input data set comprises event data.

9. (Previously Amended) A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

obtaining an input data set of items;

searching the input data set of items to identify sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and

outputting the identified mutual dependence patterns based on results of the comparisons.

10. (Canceled).

11. (Currently Amended) ~~The method of claim 10;~~ A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

obtaining an input data set of items;

normalizing the input data set;

searching the input data set of items to identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and

outputting the one or more identified patterns based on results of the comparisons,

wherein the input data set comprises event data and the normalizing step comprises transforming at least a portion of the event data into event classes such that the event data is non-application-dependent.

12. (Original) The method of claim 11, wherein the event data transformation step further comprises the step of mapping two or more attributes associated with an event into an event class.

13. (Original) The method of claim 12, wherein the mapping step is performed in accordance with a lookup table.

14. (Original) The method of claim 11, wherein the event data is in a tabular form with a first number of columns before the transformation step and in a tabular form with a second number of columns after the transformation step, the second number of columns being less than the first number of columns.

15. (Currently Amended) The method of claim ~~10~~ 11, wherein the outputting step further comprises converting the one or more identified patterns into a human readable format.

16. (Previously Amended) A computer-based method of mining one or more patterns in an input data set of items, the method comprising the steps of:

obtaining an input data set of items;

searching the input data set of items to identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and

outputting the one or more identified patterns based on results of the comparisons;

wherein the searching step further comprises the step of performing a level-wise scan based on a set length to determine candidate sets of items in the input data set that have conditional probability values respectively computed therefor that are not less than the predetermined threshold value.

17. (Original) The method of claim 16, further comprising the step of pruning candidate sets based on an upper bound property.

18. (Original) The method of claim 17, wherein the upper bound property specifies that only candidate sets are considered where the conditional probability of a set of items minus a particular subset of items given the particular subset of items is not greater than the number of occurrences of the set of items minus the particular subset of items divided by the number of occurrences of the subset of items.

19. (Previously Amended) Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) identify sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and (ii) output the identified mutual dependence patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the identified mutual dependence patterns.

20. (Previously Amended) Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and (ii) output the one or more identified patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the one or more identified patterns;

wherein the identifying operation further comprises identifying a set of items in the input data set, which includes at least two subsets of at least one item, as a pattern when the set of items has a conditional probability value computed therefor that is not less than a predetermined threshold value, wherein the conditional probability value is indicative of a probability that both of the at least two subsets of at least one item will occur given that one of the at least two subsets of at least one item has occurred.

21. (Original) The apparatus of claim 20, wherein the predetermined threshold value is between zero and one.

22. (Original) The apparatus of claim 20, wherein the conditional probability value is estimated by the number of occurrences in the input data set of the two subsets divided by the number of occurrences in the data set of one of the two subsets.

23. (Previously Amended) Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and (ii) output the one or more identified patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the one or more identified patterns;

wherein the identifying operation further comprises identifying a set of items in the input data set as a pattern when the set of items has a conditional probability value computed for the set of items minus a particular item of the set, given the particular item of the set, that is not less than a predetermined threshold value.

24. (Original) The apparatus of claim 23, wherein the predetermined threshold value is between zero and one.

25. (Previously Amended) The apparatus of claim 20, wherein the input data set comprises transaction data.

26. (Previously Amended) The apparatus of claim 20, wherein the input data set comprises event data.

27. (Previously Amended) Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) obtain an input data set of items; (ii) search the input data set of items to identify sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and (iii) output the identified mutual dependence patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the identified mutual dependence patterns.

28. (Canceled).

29. (Currently Amended) ~~The apparatus of claim 28;~~ Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) obtain an input data set of items; (ii) search the input data set of items to identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and (iii) output the one or more identified patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the one or more identified patterns;

wherein the at least one processor is further operative to, prior to the searching operation, normalize the input data set, and further wherein the input data set comprises event data and the normalizing operation comprises transforming at least a portion of the event data into event classes such that the event data is non-application-dependent.

30. (Original) The apparatus of claim 29, wherein the event data transformation operation further comprises mapping two or more attributes associated with an event into an event class.

31. (Original) The apparatus of claim 30, wherein the mapping operation is performed in accordance with a lookup table.

32. (Original) The apparatus of claim 29, wherein the event data is in a tabular form with a first number of columns before the transformation operation and in a tabular form with a second

number of columns after the transformation operation, the second number of columns being less than the first number of columns.

33. (Currently Amended) The apparatus of claim ~~28~~ 29, wherein the outputting operation further comprises converting the one or more identified patterns into a human readable format.

34. (Previously Amended) Apparatus for mining one or more patterns in an input data set of items, the apparatus comprising:

at least one processor operative to: (i) obtain an input data set of items; (ii) search the input data set of items to identify one or more sets of items in the input data set as one or more patterns based on respective comparisons of conditional probability values associated with each of the one or more sets of items to a predetermined threshold value; and (iii) output the one or more identified patterns based on results of the comparisons; and

a memory, coupled to the at least one processor, which stores at least one of the input data set and the one or more identified patterns;

wherein the searching operation further comprises performing a level-wise scan based on a set length to determine candidate sets of items in the input data set that have conditional probability values respectively computed therefor that are not less than the predetermined threshold value.

35. (Original) The apparatus of claim 34, wherein the at least one processor is further operative to prune candidate sets based on an upper bound property.

36. (Original) The apparatus of claim 35, wherein the upper bound property specifies that only candidate sets are considered where the conditional probability of a set of items minus a particular subset of items given the particular subset of items is not greater than the number of occurrences of the set of items minus the particular subset of items divided by the number of occurrences of the subset of items.



37. (Previously Amended) An article of manufacture for mining one or more patterns in an input data set of items, the article comprising a machine readable medium containing one or more programs which when executed implement the steps of:

identifying sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and

outputting the identified mutual dependence patterns based on results of the comparisons.

38. (Previously Amended) An article of manufacture for mining one or more patterns in an input data set of items, the article comprising a machine readable medium containing one or more programs which when executed implement the steps of:

obtaining an input data set of items;

searching the input data set of items to identify sets of items in the input data set as mutual dependence patterns based on respective comparisons of conditional probability values associated with each of the sets of items to a predetermined mutual dependence threshold value; and

outputting the identified mutual dependence patterns based on results of the comparisons.